Executive Summary of the Standards and Interoperability (S&I) Framework
Data Provenance Initiative Use Case Document

Introduction

In April 2014, the Office of the National Coordinator for Health Information Technology (ONC) launched the Data Provenance (DPROV) Initiative (as part of the Standards and Interoperability (S&I) Framework) to identify the standards necessary to capture and exchange provenance data, including provenance at time of creation, modification, and time of exchange.

The term "provenance" in the context of health IT refers to the evidence and attributes describing health data's origin as it is created or modified during care delivery, exchange, and other uses. The requirements for data provenance information must support the full lifecycle and lifespan of health data. As the exchange of health data increases, so does the demand to track the provenance of this data over time and with each exchange instance. Confidence in the authenticity, trustworthiness and reliability of the data being shared is fundamental to robust, privacy, safety, and security enhanced health information exchange.

The DPROV initiative recently completed a Use Case document. The Use Case is the foundation for identifying and specifying required standards as well as for developing reference implementations and tools to ensure consistent implementation according to the identified standards. The Use Case is not intended to cover every possibility or edge case, but should be considered as the basis for development of a broadly applicable technical specification which will be further refined and tested with pilots.

The Use Case document describes:

- User stories which link the goals and needs of stakeholders to functional capabilities;
- The operational context for the data exchange;
- The information flows that must be supported by the data exchange; and
- The types of data to be supported by the data exchange.

Scenarios

As a prerequisite, and applicable to all of the following use case scenarios, the Data Provenance Initiative Use Case has also defined an initial provenance metadata set. This initial set will be refined during the Initiative's "harmonization" phase, which may include adding or removing provenance metadata elements. The provenance metadata will be further categorized and described in terms of Who, When, Where, Type (What), Why, and any other information necessary to more fully describe the provenance events associated with each technical actor in the Use Case.

The Use Case consists of 3 distinct scenarios as outlined below. Note that the "start point" in each scenario could be performed by a number of different technical actors, including:

- The system which created the health data (e.g. a medical device, EHR, PHR)
- A system (such as an HIE) containing health data obtained from multiple sources
- A system (such as an EHR system) containing a combination of health data created locally and obtained from third parties

Scenario 1: Start Point → End Point. This scenario describes simple provenance requirements when transferring healthcare data from a Start Point (sending system) to an End Point (Receiving System).

Scenario 2: Start Point → Transmitter → End Point. This scenario includes use of a third party as a conduit/transmitter to transfer information from Start Point to End Point. There may be use cases where it is important to know how the information was routed, as well as who originated it and who sent it.

Scenario 3: Start Point → Assembler / Composer→ End Point. This scenario uses a third party system to aggregate or combine information from multiple sources, either in whole or in part, to produce new healthcare artifacts. The new artifacts may contain information previously obtained from multiple sources, as well as new information created locally.

Risks/Issues/Obstacles

There are several risks, issues, and obstacles associated with the management of provenance data that will need to be addressed, including management of potentially accumulated provenance information, questions about the appropriate level of granularity / format of the provenance information, ability to preserve provenance data as technology changes, and variability in system behavior which affects how provenance information is used.

Candidate Standards

The Candidate Standards List contains all standards suggested by the S&I Community that may be potentially relevant to the solution plan of the S&I Framework's Data Provenance Initiative. The candidate standards are organized into four categories:

1. **Content & Structure**: Standards defining the specific structure, content, and message format requirements for a particular Use Case
2. **Terminology & Code Values**: Standards providing specific language and naming conventions for representing clinical and medical information
3. **Transport**: Standards defining the protocols for effective data transmission to disparate systems
4. **Security**: Standards defining protocols for safe and secure transport of transmitted data

The Candidate Standards List in Table 1 shows the initial candidate standards for each of the four categories:

Content and Structure	Terminology and Code Values - DP	Transport	Security
* C-CDA * CDA R2 * HL7 Data Provenance IG CDA R2 (Draft) * W3C PROV: PROV-AQ, PROV-CONSTRAINTS, PROV-XML * Personal Health Record System Functional Model * HL7 V.2.x * HL7 EHR Records Management and Evidentiary Support (RM-ES) Functional Model, Release 1 * HL7 EHR System Functional Model Release 2 * HL7 FHIR DSTU Release 1.1 Provenance Resource * HL7 EHR Lifecycle Model (2008) * HL7 Record Lifecycle Event Metadata using FHIR (project underway 2014) * ISO 21089 Health Informatics: Trusted End-to-End Information Flows	* HL7 V.2 & V3 Vocabulary & Terminology Standards	* Representation State Transfer (RESTful) * Cross Enterprise Document-Sharing XDS * SOAP	* HL7 Implementation Guide: Data Segmentation for Privacy (DS4P), Release 1 * HL7 Health Care Privacy and Security Classification System, Release 1 * HL7 Version 3 Standard: Privacy, Access and Security Services (PASS) * HL7 Digital Signature

Table 1 - List of Candidate Standards